

多选题算法

目录

SPSSAU 数据格式..... 1
 计算公式..... 2
 1. 整体说明..... 2
 2. 计算公式..... 2
 参考文献..... 2

SPSSAU 问卷研究模块里面的‘多选题’，其统计各个选项的‘选择’频率情况，在 SPSSAU 中支持：

- ✓ 设置‘被选中’的标准，默认以数字 1 表示‘被选中’；
- ✓ 提供普及率和响应率指标；
- ✓ 提供卡方拟合优度检验。



SPSSAU 数据格式

multiple choice_option1	multiple choice_option2	multiple choice_option3
1	1	1
1	1	0
0	0	0
1	1	1
1	1	0
0	0	0
0	1	1
1	1	0
1	0	1
1	1	1
1	1	0
1	1	1
0	0	0

通常多选题的数据格式为‘1列为1个选项’多选题有多个选项，则有多个标题，并且通常情况下，数字1表示选中，数字0表示没有选中。

计算公式

1. 整体说明

SPSSAU 问卷研究模块中的‘多选题’，共用于统计多选题各个选项的选择频率，并且计算各选项选中频率的分布是否具有明显的差异性（使用卡方拟合优度检验）。

2. 计算公式

假设有 n 个样本和 k 个多选题选项。每个选项 j ($j = 1, 2, \dots, k$) 的选中状态用参数值表示，设定为：

$$X_{ij} = 1, \text{ 如果样本 } i \text{ 选择了选项 } j$$

X_{ij} 为样本 i 对选项 j 的选择状态。

提示：SPSSAU 中默认数字1表示‘选中’状态，也可通过‘计数项’参数将其设置为数字2和数字0。

每个选项 j 的频率 f_j 可以通过以下公式计算：

$$f_j = \frac{\sum_{i=1}^n X_{ij}}{n}$$

普及率（Prevalence Rate）计算公式如下：

$$\text{Prevalence Rate} = \frac{\sum_{i=1}^n X_{ij}}{n} \times 100\%$$

这里 $\sum_{i=1}^n X_{ij}$ 表示选择选项 j 的样本数量，普及率加和后正常情况下会大于100%。

响应率（Response Rate）计算公式如下：

$$\text{Response Rate} = \frac{\sum_{i=1}^n X_{ij}}{\sum_i \sum_j X_{ij}} \times 100\%$$

其中 R_i 表示样本 i 是否选中（1表示选中，0表示无效回应），响应率加和一定为100%。

关于卡方拟合优度检验，其检查普及率是否有着明显差异（假定是各选项的普及率相等），具体计算原理等可见 SPSSAU 医学研究模块里面的‘卡方拟合优度检验’。

参考文献

【1】 The SPSSAU project (2024). SPSSAU. (Version 24.0) [Online Application Software]. Retrieved from <https://www.spssau.com>.

【2】 周俊,马世澎. SPSSAU 科研数据分析方法与应用.第1版[M]. 电子工业出版社,2024.